EXAM TECHNIQUES OF ARTIFICIAL INTELLIGENCE

Thursday January 19, 14.00-17.00

- Please answer each question on a separate sheet and please put your name on all sheets.
- Work individually.
- Provide clear explanation and motivations (e.g. 5 lines, not 2 pages).
- It is an open book exam.
- Good luck!

Question 1.

Consider the problem learning to predict Z based on A,B and C. We are provided 12 training cases, as shown below:

Nr	A	B	C	Z
0	0	0	0	0
1	0	0	1	0
2	0	0	1	0
3	0	1	0	0
4	0	1	1	0
5	0	1	1	1
6	1	0	0	0
7	1	0	1	1
8	1	1	0	1
9	1	1	0	1
10	1	1	1	0
11	1	1	1	1

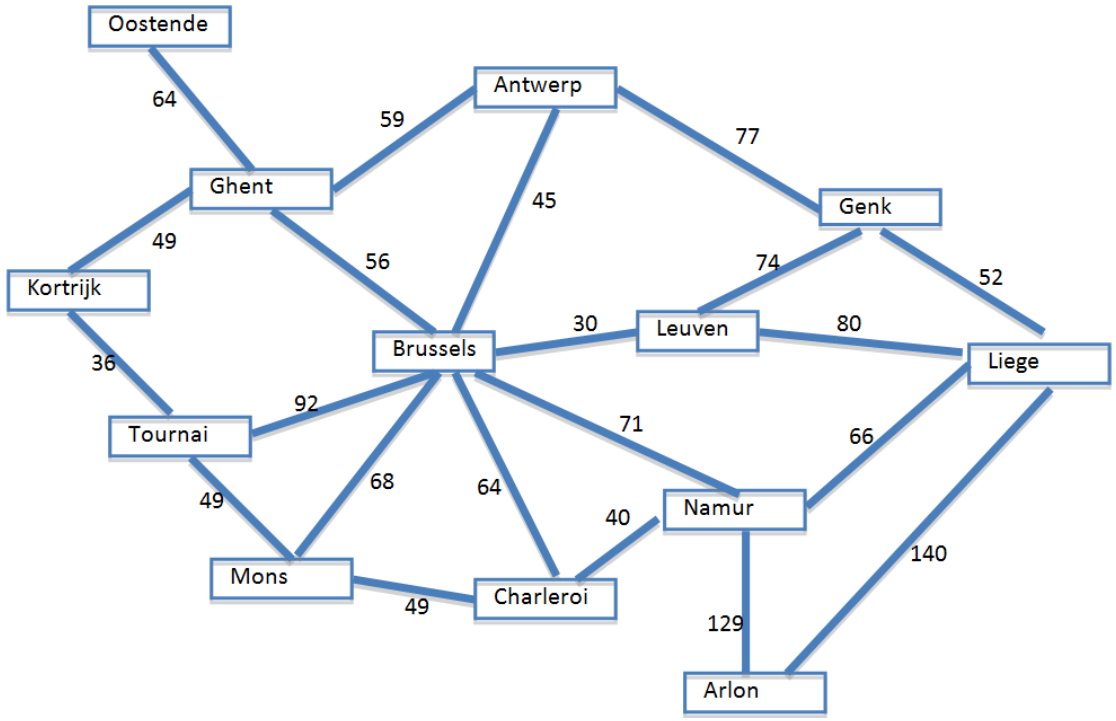
1. What will be training set error for this dataset when applying ID3? Give a motivation. (!!! This can be answered without growing the tree).
2. What will be the hypothesis learnt from Candidate Elimination? Give a motivation. (!!! This can be answered without building the version space).
3. If we apply the ID3 algorithm on this dataset, what will be topmost node of the tree? Compute the entropy of this dataset and the information gains of each attribute.
4. Apply Candidate Elimination on the instances 0,1,2,3,4. Show the version space after each training instance.
5. After training on all 12 of the above training instances, we provide a new instance for testing the learnt hypothesis . What will be the classification of the new instance A=1,B=1,C=1, if the learner is the Naïve Bayes algorithm? Show your calculations. What will be the classification of the new instance A=1,B=1,C=1, if the learner is the ID3 algorithm? Why is it that both classifications are different?

Question 2.

In back propagation, the objective function to be minimized is usually a sum of the squared error between the output and the target. However, in classification problems, the goal is often to minimize the total number of misclassifications. Why can this total number of misclassifications not be used directly as an objective function in back propagation?

Question 3.

Consider the map of Belgium, as shown below. The goal is to find the shortest path from Namur to Kortrijk.



Antwerp	91	Charleroi	96	Leuven	102	Namur	120
Arlon	222	Genk	98	Liege	165	Oostende	50
Brussels	77	Ghent	49	Mons	64	Tournai	36

Table 1. Flying distances from Kortrijk

1. Which search algorithms can be used? Which one(s) is (are) guaranteed to find the shortest path?
2. Let's assume we have, on top of the driving distances on the map, the flying distance D between Kortrijk and all other cities (see table 1). Which other algorithm(s) can then be used? Which one(s) is (are) guaranteed to find the shortest path?
3. Apply A* to find a path. Is the path optimal?
4. Is A* guaranteed to give an optimal solution using the flying distance *D* as a heuristic? With *D*/2? With 2*D*? What if there is some roadwork between Kortrijk and Tournai, increasing the distance between the two cities to 97 km?

Question 4.1.

Is reinforcement learning a subdomain of supervised learning, of unsupervised learning or is it none of both? Explain.

Question 4.2.

Consider the reinforcement learning problem given below (the numbers of the arrows represent the immediate reward), where the discount factor is $\gamma = 0.5$. Compute the V*(s) and Q*(s,a) values for each state s and show the optimal policy. Would you obtain the same V and Q values if the discount factor is set to 1?